TITLE:

<Bioinformatics Center> Bioinformatics Training Unit

AUTHOR(S):

CITATION:

<Bioinformatics Center> Bioinformatics Training Unit. ICR Annual Report 2007, 13: 66-67

ISSUE DATE:

2007-03

URL:

http://hdl.handle.net/2433/65516

RIGHT:

京都大学学術情報リポジトリ
KURENAI 紅
Kyoto University Research Information Repository

Bioinformatics Center
- Bioinformatics Training Unit -

http://www.bic.kyoto-u.ac.jp/toh/index_E.html

Vis Assoc Prof
KUMA, Kei-ichi
(D Sc)

Vis Assist Prof
ICHIHARA, Hisako

Student

SATO, Tetsuya (D3)

Scope of Research

Evolutionary studies based on molecular biology is called "molecular evolutionary biology", which is one of the origins of the current bioinformatics. Living organisms have acquired wide variety of functions during the course of the evolution by changing the information encoded by the genomes. Inversely, reconstruction of the evolutionary history related to the functions would bring us a great insight into the acquired functions and the life. Furthermore, such evolutionary information is useful for practical fields such as drug design and proteins engineering. We develop new methodologies with evolutionary information, to extract biological knowledge from various molecular biological data including sequence and structure data of individual genes and proteins, genome data, and expression profile data. We also analyze the data of molecular biology from the evolutionary viewpoint, to obtain novel biological knowledge.

Research Activities (Year 2006)

Presentations

Evolutionary Analysis of Proteins Relevant to Quorum Sensing, Ichihara H, Kuma K, Toh H (Medical Institute of Bioregulation, Kyushu University), 20th IUBMB Int'l Congress of Biochemistry and Molecular Biology and 11th FAOBMB Congress, Japan, 21 June.

A Study of Relationship between Mammalian Specific Features and Gene Diversification on the Basis of Genome Comparisons, Kuma K, 7th Society of Evolutionary Studies, Japan, 30 August.

Positive Selection in CSP-ComD System of Streptococcal Species, Ichihara H, Kuma K, Toh H (MiB, Kyushu Univ.), MBSJ 2006 Forum, Japan, 7 December (Oral Presentation).

Relationship between Data Size and Accuracy of Prediction of Protein-Protein Interactions by Co-evolutionary Information, Sato T, Yamanishi Y, Ichihara H, Kuma K, Kanehisa M, Toh H (MiB, Kyushu Univ.), 17th Int'l Conference on Genome Informatics, Japan, 18–20 December.

Grant

Kuma K, A Study of Relationship between Mammalian Specific Features and Gene Diversification on the Basis of Genome Comparisons, Grant-in-Aid for Scientific Research (C), April 2005–March 2007.

Partial Correlation Coefficient between Distance Matrices as a New Indicator of Protein-Protein Interactions

The computational prediction of protein–protein interactions is currently a major issue in bioinformatics. Recently, a variety of co-evolution-based methods have been investigated toward this goal. In this study, we introduced a partial correlation coefficient as a new measure for the degree of co-evolution between proteins, and proposed its use to predict protein–protein interactions.

The accuracy of the prediction by the proposed method was compared with those of the original mirror tree method and the projection method previously developed by our group. We found that the partial correlation coefficient effectively reduces the number of false positives, as compared with other methods, although the number of false negatives increased in the prediction by the partial correlation coefficient. The R script for the prediction of protein–protein interactions reported in this manuscript is available at **http://timpani.genome.ad.jp/~parco/**

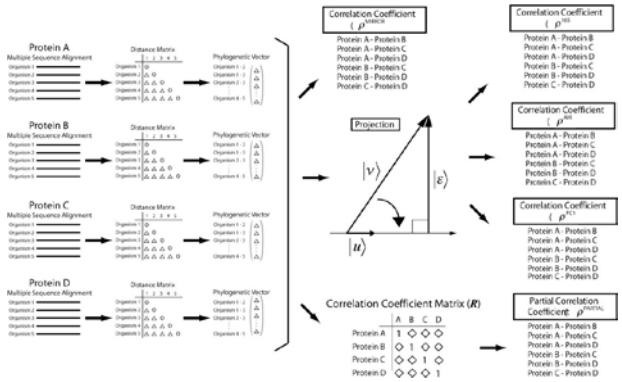

Figure 1. A schematic representation of the procedures for predicting interaction partners.

Adaptive Evolution in the ComC-ComD System of Streptococcal Species

ComC and ComD of the streptococcal species are a peptide pheromone and its receptor, respectively, which are involved in the regulation of natural DNA uptake from the environment. Both ComC and ComD show high sequence variability. Recombination is regarded as a mechanism to generate the sequence diversity. In this study, we investigated another possibility for generating sequence diversity, adaptive point mutation, after excluding the effect of recombination as much as possible. We evaluated the rate of the number of non-synonymous substitutions per site to the number of synonymous substitutions per site by two different methods. The results of both approaches suggested that the C-terminal region of ComC and the N-terminal region of ComD have undergone positive selection. In addition, the binding specificity of ComC to the cognate ComD suggests that the adaptive change may have occurred through co-evolution between ComC and ComD. The meaning of the positive selection in the ComC-ComD system is discussed from the viewpoint of competition among the streptococcal strains for DNA resources.

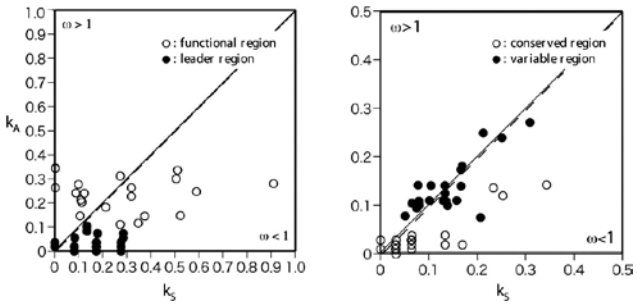

Figure 2. The k_A values are plotted as a function of the k_S values. If a plot is present above the diagonal dashed line, then the omega value for the pair corresponding to the plot is >1.0.

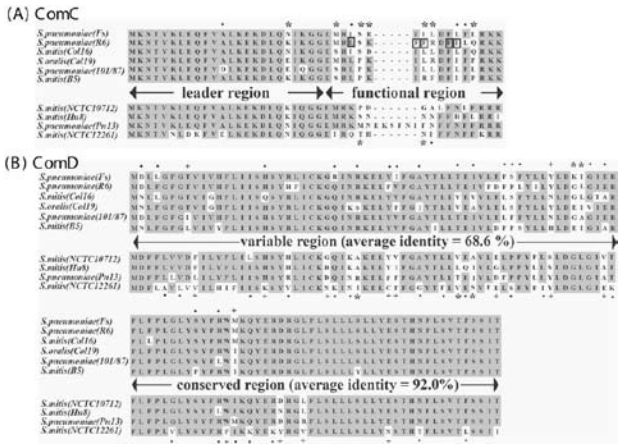

Figure 3. Multiple alignments of the ComC (A) and ComD (B) amino acid sequences. The asterisks, plus signs, and period symbols indicate the amino acid sites that correspond to the codon sites that have an omega of >1.0, with posterior probabilities higher than 99, 95, and 50%, respectively.